

# Anticipating critical transitions in multi-dimensional systems driven by time- and state-dependent noise

Andreas Morr,<sup>1,2,\*</sup> Keno Riechers,<sup>1,2</sup> Leonardo Rydin Gorjão,<sup>3</sup> and Niklas Boers<sup>1,2,4</sup>

<sup>1</sup>*Earth System Modelling, School of Engineering and Design, Technical University of Munich, Germany*

<sup>2</sup>*Research Domain IV – Complexity Science, Potsdam Institute for Climate Impact Research, Germany*

<sup>3</sup>*Faculty of Science and Technology, Norwegian University of Life Sciences, Norway*

<sup>4</sup>*Global Systems Institute and Department for Mathematics, University of Exeter, UK*

(Dated: September 1, 2023)

The anticipation of bifurcation-induced transitions in dynamical systems has gained relevance in various fields of the natural, social, and economic sciences. When approaching a co-dimension 1 bifurcation, the feedbacks that stabilise the initial state weaken and eventually vanish; a process referred to as critical slowing down (CSD). This motivates the use of variance and lag-1 autocorrelation as indicators of CSD. Both indicators rely on linearising the system’s restoring rate. Additionally, the use of variance is limited to time- and state-independent driving noise, strongly constraining the generality of CSD. Here, we propose a data-driven approach based on deriving a Langevin equation to detect local stability changes and anticipate bifurcation-induced transitions in systems with generally time- and state-dependent noise. Our approach substantially generalizes the conditions underlying existing early warning indicators, which we showcase in different examples. Changes in deterministic dynamics can be clearly discriminated from changes in the driving noise. This reduces the risk of false and missed alarms of conventional CSD indicators significantly in settings with time-dependent or multiplicative noise. In multi-dimensional systems, our method can greatly advance the understanding of the coupling between system components and can avoid risks of missing CSD due to dimension reduction, which existing approaches suffer from.

## I. INTRODUCTION

A mechanistic understanding of complex high-dimensional physical systems is essential for assessing the risk of abrupt regime shifts, for example in ecological, climatic, social, or financial systems. Such shifts may occur when critical forcing thresholds, which correspond to underlying bifurcation points, are crossed [1–4]. Reducing complex systems to a low-dimensional summary observable  $\mathbf{X}_t$  has leveraged impressive modelling capabilities [5–8]. This is particularly important because observations are typically available in the form of multivariate time series of just a few dimensions.

Commonly, the dynamics of the summary observable  $\mathbf{X}_t \in \mathbb{R}^n$  is approximately separated into a deterministic component  $A(\mathbf{X}_t, t)dt$  and a stochastic component  $B(\mathbf{X}_t, t)d\mathbf{W}_t$  that represents the action of the omitted dimensions. This results in the Langevin equation

$$d\mathbf{X}_t = A(\mathbf{X}_t, t)dt + B(\mathbf{X}_t, t)d\mathbf{W}_t. \quad (1)$$

Even though, in principle, the stochastic component can take more complicated forms, we restrict ourselves to the case where  $\mathbf{W}$  is an uncorrelated Wiener process supported on the filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}_+}, \mathbb{P})$  and refer to existing extensions to the case of correlated noise [9, 10].

This framework facilitates a mathematical description of abrupt regime shifts in terms of dynamic bifurcations in low-dimensional dynamical systems [2, 11]. Prior to

the transition, the deterministic drift  $A(\mathbf{X}_t, t)$  embodies negative feedback mechanisms keeping the system in a stable equilibrium [12–14]. The explicit time dependence of  $A(\mathbf{X}_t, t)$  reflects the changing forcing levels that act on the system from the outside and alter the deterministic, coarse-grained dynamics. At the bifurcation point, i.e. at the critical level of forcing, the currently occupied equilibrium state is annihilated and the system abruptly transitions to another stable state.

For co-dimension 1 bifurcations, it is well known that a weakening of the negative feedback precedes an eventual abrupt transition [2, 15]. This will, heuristically speaking, result in a weaker and slower response to the pseudo-random perturbations stemming from the unresolved dynamics. This phenomenon is referred to as critical slowing down (CSD), and it manifests in an increase of the statistical quantities of variance and lag-1 autocorrelation (AC(1)) of the observable in the components exhibiting stability loss [2, 16–18]. These two quantities are therefore often employed to anticipate bifurcation-induced abrupt transitions, and their simultaneous increase has been suggested as an early warning signal (EWS) [2, 3, 19]. Mathematically, CSD can be described by approximating the negative feedback around a stable equilibrium as a linear restoring rate. Denoting the time-dependent equilibrium state of a one-dimensional observable  $X_t$  by  $x^*(t)$ , we arrive at the Ornstein–Uhlenbeck model [20]

$$dX_t = -\lambda(t)(X_t - x^*(t))dt + \sigma dW_t.$$

The linearised negative feedback  $\lambda(t)$  weakens during CSD while the noise-coupling strength  $\sigma$  is assumed to remain constant. This results in the following expressions

---

\* [andreas.morr@tum.de](mailto:andreas.morr@tum.de)

for variance and AC(1),

$$\begin{aligned} \text{Var}[X] &= \frac{\sigma^2}{2\lambda} \xrightarrow{\lambda \rightarrow 0} \infty \\ \text{AC}_X(1) &= \exp(-\lambda\Delta t) \xrightarrow{\lambda \rightarrow 0} 1, \end{aligned} \quad (2)$$

where  $\Delta t > 0$  is the sampling time step of the data. Detection of CSD is usually preceded by a reduction of the system to a one-dimensional observable, either by leveraging physical understanding or employing principle component analysis [16, 21] in order to identify a linear combination of components which may be experiencing stability loss. In such a reduction, crucial information about system stability may be lost. The method presented herein is applicable directly to data from higher dimensional systems (or to multivariate data) and thus avoids this preprocessing step.

We will nevertheless, for illustrative purposes, first treat the problem of estimating local system stability in the one-dimensional dynamical system denoted by

$$dX_t = a(X_t, t)dt + b(X_t, t)dW_t, \quad (3)$$

such that the linearised negative feedback takes the form

$$\lambda(t) := -\partial_x a(x^*(t), t). \quad (4)$$

The extension of the discussed estimation methods to the general, higher-dimensional setting of Eq. (1) is discussed in the Methods section.

Time- or state-dependent driving noise can lead to both false negative and false positive EWS [9, 22, 23]. Therefore, understanding the evolution of the diffusion term  $b(X_t, t)dW_t$  is crucial for reliable statements on stability changes derived from data. Given that in real-world situations, the assumption of time- and state-independent noise is hardly justifiable, a more general theoretical framework advancing CSD to the case of time- and state-dependent noise is called for.

In particular, a methodology is needed to extract from the observable a more holistic picture of both the deterministic dynamics of the system and the driving noise. The derivation of the variance and AC(1) in (2) hinges on the a priori assumption of linear feedback. In applications, the system might explore parts of the state space where non-linearities in the feedback are not negligible anymore, putting the validity of Eq. (2) into question. In contrast, the linear restoring rate  $\lambda$  directly captures the desired information of local stability and should therefore be considered the key quantity to measure system stability and detect CSD. To obtain an estimation  $\hat{\lambda}$ , we perform a spatially local linear fit to the estimated function  $\hat{a}(x)$  [24, 25] in some neighbourhood around  $\hat{x}^* = \hat{\mathbb{E}}[X]$  (see Methods and Supplementary Material (SM) 1). A similar approach has recently been proposed in [26]. We carry the concept to multiple dimensions and include an estimation of the diffusion matrix  $BB^\top(x, t)$  to supplement the standard CSD indicators and to avoid false positives and false negatives caused by changes in the driving

noise. In particular, we discuss situations where the conventional CSD indicators give an ambiguous or misleading picture. We show how the method proposed herein conclusively resolves these ambiguities.

## II. METHODS

It can be shown that the drift and diffusion coefficients  $A$  and  $B$  have the following representation in terms of the increments  $\Delta\mathbf{X}_t := \mathbf{X}_{t+\Delta t} - \mathbf{X}_t$  of the process  $\mathbf{X}$  [24, 27–33]:

$$\begin{aligned} A(\mathbf{x}, t) &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \mathbb{E}[\Delta\mathbf{X}_t | \mathbf{X}_t = \mathbf{x}], \\ BB^\top(\mathbf{x}, t) &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \mathbb{E}[\Delta\mathbf{X}_t \Delta\mathbf{X}_t^\top | \mathbf{X}_t = \mathbf{x}]. \end{aligned}$$

If the stochastic differential equation (SDE) (3) exhibits effective time independence, i.e.  $A(\mathbf{x}, t) \equiv A(\mathbf{x})$  and  $B(\mathbf{x}, t) \equiv B(\mathbf{x})$  in some observation time span and if the sample path of  $\mathbf{X}$  is available at sufficiently small time steps  $\Delta t > 0$ , one may estimate  $A(\mathbf{x})$  and  $BB^\top(\mathbf{x})$  by replacing the above ensemble average by the mean of the observed increments. The law of large numbers yields consistent estimators that converge to the true  $A$  and  $BB^\top$ , omitting here a small bias stemming from the non-zero  $\Delta t$ .

We generalise the definition of local system stability  $\lambda$  in Eq. (4) from the one-dimensional setting to the multi-dimensional setting of Eq. (1). Consider the Jacobian matrix of  $A$  at an equilibrium point  $\mathbf{x}^* \in \mathbb{R}^n$  with  $A(\mathbf{x}^*, t) = \mathbf{0}$ . Such an equilibrium point is stable if and only if all eigenvalues  $(-\lambda_k + i\omega_k)_{k=1, \dots, n}$  of the Jacobian matrix  $DA(\mathbf{x}^*, t)$  exhibit negative feedback  $-\lambda_k < 0$ . Accordingly, we regard the set of  $(\lambda_k)_{k=1, \dots, n}$  as a measure of system stability.

We obtain an estimation of  $A$  and  $BB^\top$  from time series data using the respective estimators given in [25, 34]. A multivariate ordinary least squares regression is performed to extract an estimate of the matrix  $DA$  (see SM 1 for details). The real parts of the corresponding eigenvalues are then assessed in their time evolution. In a windowed time series analysis, negative trends in any of the  $\lambda_k$  would indicate a destabilisation of the equilibrium state, which may point to an upcoming abrupt transition. The windows must be short enough to justify the required time independence of the dynamics within each individual window and yet comprise a sufficient amount of data. The windowed estimation of  $A(\mathbf{x}, t)$  and  $BB^\top(\mathbf{x}, t)$  then reveals potential temporal changes in the system's stability.

We will show that the estimation of the diffusion matrix  $BB^\top$  may in certain situations explain peculiar behaviour in the conventional CSD indicators, such as simultaneously decreasing variance and increasing AC1. If instead of local stability, here measured by  $\lambda$ , one wishes to examine mean exit times from stable equilibria [35],

the estimated diffusion matrix  $BB^\top$  has additional important implications [36].

### III. RESULTS

First, we will examine the merits of the proposed CSD indicator  $\hat{\lambda}$  on the estimated drift coefficient in a conceptual, one-dimensional system subjected to time-dependent noise. Second, we will adopt a two-dimensional predator-prey model with state-dependent noise from the literature and perform a multi-dimensional analysis.

#### A. Fold bifurcation with time-dependent noise

The prototypical structure employed for conceptualising abrupt transitions in many natural systems is the fold bifurcation [2, 37]. Consider therefore the SDE defined by

$$dX_t = (-X_t^2 + \alpha(t))dt + \sigma(t)dW_t, \quad (5)$$

where  $\alpha$  is the bifurcation parameter and  $\sigma$  the noise strength. For positive  $\alpha > 0$ , there exists a stable equilibrium at  $x^*(t) = \sqrt{\alpha}$  which vanishes at the critical threshold  $\alpha_{\text{crit}} = 0$  (see Fig. 1a). We simultaneously ramp down the noise strength  $\sigma(t)$ . Such an evolution should be understood as a change in the nature of the omitted fast dynamics, which cannot be ruled out in many applications [38, 39].

The temporal evolution of the variance and AC(1) can be approximated by Eq. (2) after linearising the system around the time-dependent equilibrium  $x^*(t)$  (red lines in Fig. 1b and c):

$$\text{Var}[X_t] \approx \frac{\sigma(t)^2}{4\sqrt{\alpha(t)}}, \quad \text{AC}_X(1) \approx \exp\left(-2\sqrt{\alpha(t)}\Delta t\right).$$

The time-dependent noise thus induces a deceiving downward trend in the estimated variance of the system (Fig. 1b) alongside increasing AC(1) (Fig. 1c). The conflicting indications given by variance and AC(1) would mislead the observer to conclude that no significant EWS is present. In contrast, the estimation of the linearised feedback  $\lambda$  (Fig. 1d) clearly indicates a weakening of local system stability and thus the presence of CSD.

The apparently inconsistent results of the conventional CSD indicators can be understood and reconciled by examining the structure of the drift and diffusion coefficients  $a(x,t) = x^2 - \alpha(t)$  and  $b(x,t) = \sigma(t)$  and their evolution in time. The true quantities for the functions  $a(x,t)$  and  $b(x,t)$  at different times  $t$  are plotted in Fig. S1 in SM 1 along with the estimations obtained during the procedure outlined in the methods section. After disclosing the time dependency of the diffusion coefficient in Fig. S1c, the diverging trends in variance and AC(1) can be correctly interpreted in a CSD assessment. The decrease in variance can be attributed to the decreasing

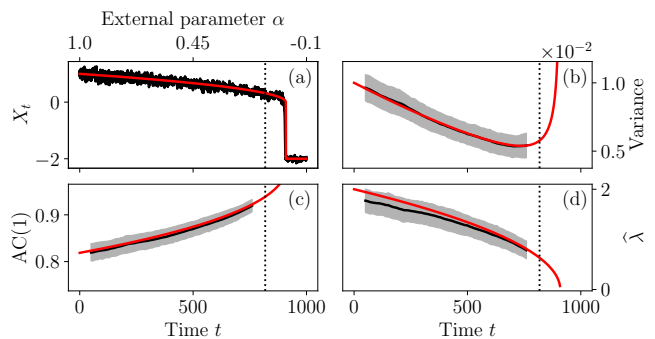


FIG. 1. Application of CSD indicators for synthetic data generated by the model (5) in the main text. (a) Sample paths for zero noise (**red**) and with noise (**black**). The noise strength  $\sigma$  is ramped linearly from 0.2 to 0.06 over the integration time span. (b), (c) Conventional CSD indicators of variance and lag-1 autocorrelations (**black**) are calculated on detrended windows and plotted along with the theoretical values (**red**) obtained from the time-local Ornstein–Uhlenbeck linearisation. The shaded bands represent the 68% confidence intervals on  $N = 1000$  samples. (d) Estimator  $\hat{\lambda}$  as proposed in this work, along with the true value  $\lambda = -\partial_x a(x^*(t), t)$  (**red**). All estimations were performed on running windows of length  $10^2$  consisting of  $10^3$  data points, considering the sample time-step is  $\Delta t = 10^{-1}$ . A traditional analysis using the variance and AC(1) would lead to a missed alarm, given their opposing trends; in contrast, the CSD methodology proposed here clearly detects the forthcoming bifurcation and correctly attributes the negative variance trend to the decreasing amplitude of the driving noise  $\sigma(t)$ .

noise strength, and the approach of a bifurcation can be confirmed. In an analogous setting featuring no bifurcation but an increasing noise strength, the incurred increase in variance can be attributed correctly and the false alarm that a conventional CSD analysis would raise can be avoided.

The statistical quality of the estimator  $\hat{\lambda}$  with respect to its distribution width is similar to that of the conventional indicators and sufficiently good to ensure a high likelihood of a statistically significant trend. This can be argued by checking that the confidence intervals at the beginning and the end of the estimator time series do not overlap (see also SM 2).

#### B. Predator-prey model with state-dependent noise

Following the work of Bengfort et al. [40], we examine a predator-prey model for oceanic plankton populations (see SM 3 for details). Since Bengfort et al. consider this model under the assumption of no external disturbances in the form of noise, we adopt a noise model from a related study [41]. Because environmental variability usually does not influence the population sizes directly, but rather their growth rates, a multiplicative noise term is

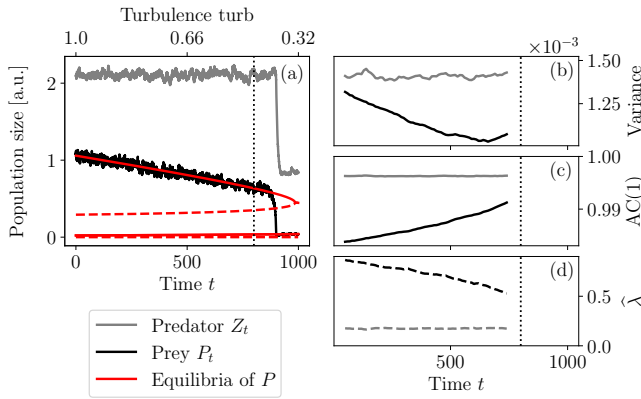


FIG. 2. Application of the CSD indicators on time-series data obtained from the predator-prey model of [40]. (a) Sample paths of the prey population  $P$  (black) and the predator population  $Z$  (grey). The stable and unstable equilibria of  $P$  in their dependence on  $\text{turb}(t)$  are plotted in red. (b) and (c) show the means of the conventional CSD indicators variance and  $\text{AC}(1)$  over  $N = 1000$  samples. (d) Real parts  $-\lambda_{1,2}$  corresponding to the estimated eigenvalues of the local Jacobian matrix. The eigenvalues have been assigned (by colour) to the two populations, as the corresponding eigenspace basis aligns very well. All estimations were performed on running windows of length 100, meaning 5000 data points at sampling rate  $\Delta t = 2 \cdot 10^{-2}$ .

often employed [42–45]. This leads us to investigate the following system of SDEs

$$\begin{aligned} dP_t &= \xi^{-1} \left( rP_t \left( 1 - \frac{P_t}{K(\text{turb})} \right) - \frac{aP_t^2}{h(\text{turb})^2 + P_t^2} Z_t \right) dt \\ &\quad + \xi^{-1/2} \sigma_P P_t dW_t^P, \\ dZ_t &= \left( \frac{aP_t^2}{h(\text{turb})^2 + P_t^2} Z_t - mZ_t^2 \right) dt + \sigma_Z Z_t dW_t^Z, \end{aligned}$$

under the external forcing of ocean turbulence. Due to the quadratic mortality term  $mZ^2$  of the predator population  $Z$ , this system can exhibit multiple stable equilibria and bifurcations as indicated in Fig. 2a. The two white noise terms are assumed to be independent and their strengths  $\sigma_P$  and  $\sigma_Z$  are chosen such that noise-induced tipping only occurs in close proximity to the bifurcation point.

Here, we examine the performance of the conventional and the newly proposed CSD indicators as the system approaches this bifurcation. Fig. 2a shows sample paths for the predator and prey populations along with the stable and unstable equilibria of the prey population  $P$  as implied by the parameter value  $\text{turb}$  at time  $t$ .

The most common approach to the assessment of CSD in a multi-dimensional system such as this one is to first reduce the system to one dimension [8, 46, 47]. The centre manifold theorem states that in close proximity to a critical bifurcation, the direction of lowest stability will be the one to experience further destabilisation. For this

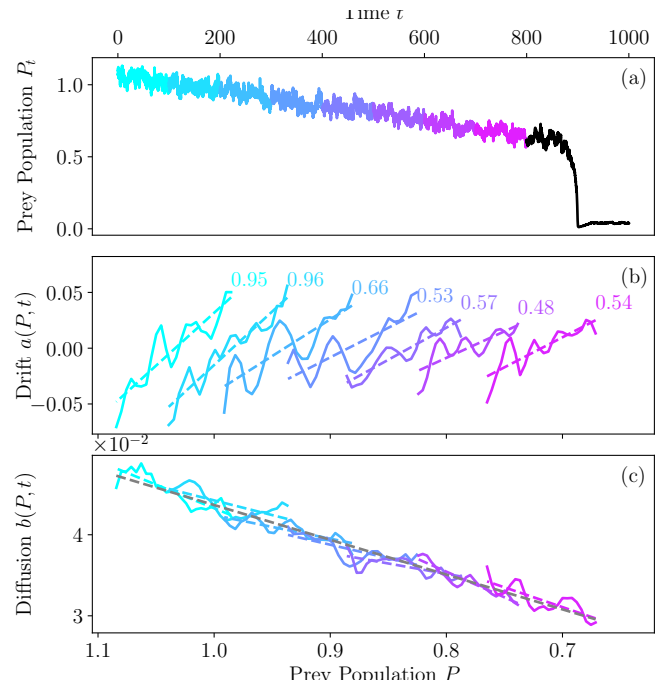


FIG. 3. Prey population taken as a one-dimensional system (a) and corresponding drift and diffusion coefficients  $a(P, t)$  (b) and  $b(P, t)$  (c) for different time slices. In (b) and (c) the functions are plotted in their  $P$ -dependence, while the time  $t$  is represented by the respective colour of the plot, as indicated in (a). The dashed lines in (b) represent the best linear fits performed on the estimated  $a(P, t)$ . Annotated is the value of the estimator  $\hat{\lambda}$  on each window of data, i.e., the negative of the slope of the respective linear fit. Similarly, linear fits in (c) are shown to illustrate the apparent state-dependence, in addition to a linear fit of the entire data (grey dashed).

reason, a principle component analysis is often performed to determine a linear combination of system components that exhibits the largest variance or  $\text{AC}(1)$  and can thus be suspected to be of the lowest stability [16, 17, 21]. However, as can be seen in the example at hand, away from the immediate proximity of the bifurcation point, the destabilising direction need not be the direction of lowest stability. Here, the identified direction of stability loss would be closely aligned with the predator population  $Z$ , as it operates on a slower time scale. This is problematic, as this dimension is relatively impervious to changes in the control parameter  $\text{turb}$  and will not exhibit CSD (grey curves in Fig. 2b and c).

To circumvent this issue, one should therefore perform a comprehensive stability analysis on the multi-dimensional time series.

This is achieved by examining both eigenvalues of the local equilibrium dynamics as shown in Fig. 2d. The real part of the larger eigenvalue can be seen to substantially decrease. This provides evidence for a destabilisation along the more stable direction in the eigenspace. Note that the conventional approach of focusing on the direction of largest eigenvalue would miss this destabi-

sation.

Investigating again the conventional variance and AC(1) CSD indicators in Fig. 2b and c, AC(1) seems to indicate a destabilisation along the dimension of the prey population  $P$  but the trend of the observed variance seems to indicate the opposite. This contradiction can be resolved by performing an analogous analysis of the drift and diffusion coefficients for the one-dimensional time series of  $P$  as in the first example above (see Fig. 3). This reduction in dimension can now be motivated by the fact that the vector in eigenspace corresponding to the weakening eigenvalue in 2d lies predominantly in the direction of  $P$ . The slopes of the estimated drift coefficient shown in Fig. 3b decrease as the system moves towards the bifurcation, agreeing with the estimations in Fig. 2d. A clear state-dependence can be identified in the estimated diffusion coefficient (Fig. 3c). Together with the observation of a diminishing mean state in the prey population, it can be concluded that the decrease in variance in 2b was due to a reduced noise amplitude and can thus be reconciled with the increase in AC(1).

#### IV. DISCUSSION

Variance and AC(1) are often used in combination to assess whether or not a system is approaching a critical transition. In general, a positive result is considered robust when both indicators show a significant positive trend. We have shown here that in the presence of time- or state-dependent noise amplitudes, the variance of the system may actually decrease in the advent of a bifurcation. If a monitored system shows a decreasing trend in variance alongside an increasing trend in AC(1), this would typically not be considered a robust EWS, leading to a missed alarm. An increase in noise strength over time could, on the other hand, lead to a false alarm in form of an increasing variance in systems with no underlying bifurcation. We have also shown that common methods in dimension reduction can lead to missed alarms, as the destabilising system component may not be the least stable to begin with.

To overcome these problems, we have proposed a method based on deriving a Langevin equation from the observed dynamics. Our approach allows us to separate the effects of possible CSD dynamics contained in the drift coefficient from changes in the noise represented by the diffusion coefficient. It also allows for a more holistic investigation of multi-dimensional systems, without further mechanistic simplifications (see SM 4 for a second example to this point, which shows in particular that the proposed method works for periodic multi-dimensional systems that are problematic for the conventional CSD indicators). We showed that our approach avoids the pitfalls that a conventional CSD analysis suffers from in these examples.

We have shown that in the presented one-dimensional application, the statistical quality of the estimator  $\hat{\lambda}$  is

of the same order as that of the estimators for variance and AC(1) (see also SM 2 for further discussion). However, one important caveat bears mentioning: While for the estimators of variance and AC(1), the length of the time series is the only determining factor of convergence, the estimators for the drift and diffusion coefficients also require small sampling time steps  $1 \gg \lambda \Delta t > 0$  in order for their bias to be small. In general, the estimator  $\hat{\lambda}$  proposed here will still contain information about CSD even in settings of large sample time steps  $\Delta t$ , but the signal-to-noise ratio may prohibit its employment as a CSD indicator. Areas of application where systems are potentially susceptible to tipping and where high-frequency data may be available for analysis could be electricity grids [48–50], financial markets [51–53], atmospheric circulation systems such as monsoons [54, 55], ecosystems and vegetation systems such as the Amazon rainforest, [56–59], ocean circulation systems [3], or ice sheets [60, 61].

Our method should be understood as a more general, reliable, and circumspect indicator of CSD compared to the widely used variance and AC(1). Our approach is appropriate in settings of generally time- and state-dependent driving noise, where the combined conventional indicators fail. Moreover, the ability to examine time series in their multi-dimensional complexity constitutes a considerable improvement in the comprehension of the system compared to one-dimensional summary statistics.

#### DATA AVAILABILITY

Supplementary Material is available for download at the online version of this manuscript. The implementation of the estimators introduced in this work is available in the GitHub repository [KramersMoyaleWS](#). Also included is the code employed to generate all figures in the main text and the Supplementary Material.

#### ACKNOWLEDGMENTS

This work has received funding from the Volkswagen Stiftung, the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 820970 and under the Marie Skłodowska-Curie grant agreement No. 956170, as well as from the Federal Ministry of Education and Research under grant No. 01LS2001A. This is TiPES contribution #X.

#### SUPPLEMENTARY MATERIAL

##### 1. Details on the estimator $\hat{\lambda}$

Here, we give a detailed description of the local stability measure  $\hat{\lambda}$  in  $n$ -dimensional systems. For  $n = 1$ ,

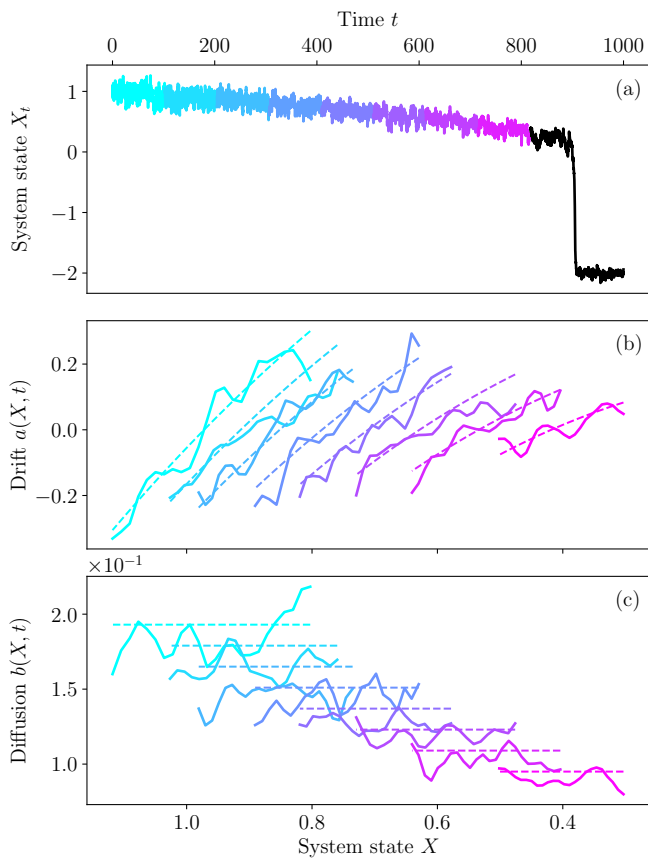


FIG. S1. Estimated drift and diffusion coefficients  $a(X, t)$  and  $b(X, t)$  [25] for the one-dimensional fold bifurcation example in the main text. The functions are plotted in their  $X$ -dependence, while the time  $t$  is represented by the respective colour of the plot. The dashed lines show the theoretical quantities.

it is the negative of the slope of the estimated drift coefficient  $a$  around the equilibrium, and therefore always a real number. For  $n > 1$ , the estimation procedure returns  $n$  eigenvalues of the local Jacobian matrix of the estimated drift coefficient in their algebraic multiplicity. The eigenvalues may be complex, and thus the value of interest investigated in the main text is the negative of the real part of the respective eigenvalues.

Prior to any analysis, the windowed time series data of each of the  $n$  dimensions is linearly detrended. For the assessment of the conventional CSD indicators, the mean of the data is removed, as they rely purely on the fluctuations around the equilibrium state. In contrast, drift and diffusion are assessed without subtraction of the mean to retain information about the corresponding state dependence. In order to obtain numerical stability, the  $n$  time series are normalised to a standard deviation of 1, with no implication on the subsequent estimations. For each window, the estimation of the function  $A(\mathbf{x})$  is returned as an array of values

$$(\hat{A}(\mathbf{x}_i))_{i=1, \dots, M^n},$$

where  $M$  is the number of evenly spaced target bins in each dimension. The  $M^n$  target bins in  $\mathbb{R}^n$ , therefore, form a grid on the hypercube spanned by the state space explored by the time series. For each of these bins, an estimation  $\hat{A}(\mathbf{x}_i)$  is calculated using an Epanechnikov kernel

$$K(\mathbf{x}) = \frac{3}{4h} \left( 1 - \frac{\|\mathbf{x}\|^2}{h^2} \right), \text{ with support } \|\mathbf{x}\| < h,$$

with a kernel bandwidth  $h$  of  $14n/M$ . Since the estimator  $\hat{A}(\mathbf{x}_i)$  will converge to some (biased) value as the number of samples  $\mathbf{X}_{k\Delta t}$  in the bin  $\mathbf{x}_i$  tends to infinity, it is clear that those bins with many samples converge fastest. In our setting with equilibrium dynamics around one stable equilibrium  $\mathbf{x}^*$ , this means that the estimations for bins closest to  $\mathbf{x}^*$  converge fastest, and the quality deteriorates for outer bins. For this reason and in order to curtail the effects of a non-linear drift term, we opt to only carry 50% of bins centred around the bin containing  $\hat{\mathbf{x}}^* = \hat{\mathbb{E}}[\mathbf{X}]$  to the subsequent analysis. This is to say, we select a hypercube with side lengths 50% as large as the original hypercube. Thus, we are confronted with fixing three free parameters a priori: The number of total bins  $M$  in each dimension, the percentage  $m$  of bins to carry on either side of the estimated equilibrium  $\hat{\mathbf{x}}^*$ , and the kernel bandwidth. In this study, we chose  $M = 50$  and  $m = 50\%$ , meaning that for  $n = 1$ , we have 25 relevant bins for further analysis. The bandwidth is chosen as a function of  $M$  and  $n$ , as described above. However, the performance of the estimator  $\hat{\lambda}$  is not very sensitive to small changes in these parameters.

To obtain an estimation of the local Jacobian matrix around the equilibrium point  $\hat{\mathbf{x}}^*$ , we perform a multivariate ordinary linear regression between  $(\hat{A}(\mathbf{x}_i))$  and  $(\mathbf{x}_i)$  over  $i$ , including an intercept in the design matrix. The algebraic eigenvalues of the resulting matrix are computed numerically. For  $n = 1$ , this procedure is equivalent to finding a best linear fit  $(c - \hat{\lambda}x_i)$  to  $(\hat{a}(x_i))$  over  $i$ .

## 2. Assessing the statistical quality of $\hat{\lambda}$

The two applications presented in the main text demonstrate that CSD manifests itself in a substantial negative trend of the estimator  $\hat{\lambda}$  when enough data are available. In this section, we aim to make this statement concrete and to compare the indicator's performance to that of variance and AC(1). The assessment is based on the width of the three indicators' numerical distribution after application to synthetic data in one dimension. The top row of Fig. S2 shows the distributions of the estimators that arise from the application of the indicators to 1000 synthetic time series generated by numerically integrating a time-homogeneous OU process:

$$dX_t = -\lambda X_t dt + dW_t, \quad (6)$$

To analyse the behaviour of the estimators in a generic CSD scenario, i.e., a temporal reduction of the restoring

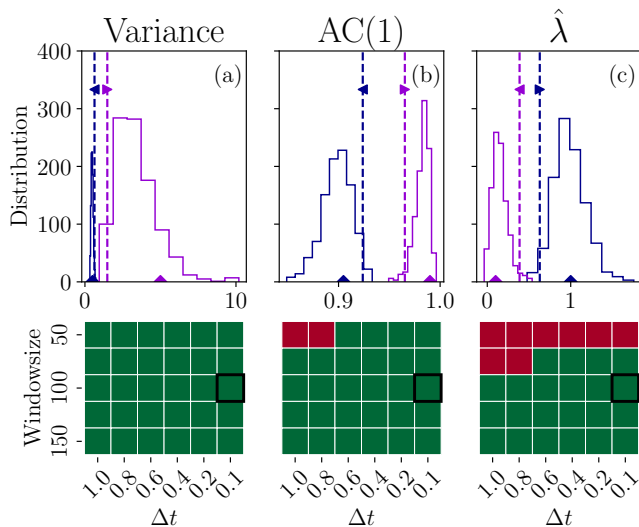


FIG. S2. Distributions of (a) variance, (b) AC(1), and (c) the estimator  $\hat{\lambda}$  on 1000 samples of Ornstein-Uhlenbeck time series of length  $T = 100$  each (see equation (6)). The sampling time step was set at  $\Delta t = 0.1$ . This estimation setting is equivalent to that chosen for the synthetic fold bifurcation example in the main text. The blue histogram shows the estimator distribution for the OU-parameter  $\lambda = 1$  and the violet for  $\lambda = 0.1$ . If the 97.5- and 2.5-percentile of the two distributions do not overlap, it is highly likely that the corresponding CSD will be detected by the indicators. The separation in distribution is investigated for several choices of  $T$  and  $\Delta t$  in the bottom row of the figure. Green indicates no significant overlap of the estimator distributions, i.e., a successful CSD estimation; red the contrary.

rate, we plot their distributions for  $\lambda = 1$  and  $\lambda = 0.1$ . If the distributions are sufficiently distinct, the indicators may correctly detect a given reduction of the restoring rate with a high likelihood. For different choices of window lengths and time steps  $\Delta t$ , we check numerically whether this condition is satisfied for each estimator (bottom row of Fig. S2). Being more sensitive at low data availability, the estimators for variance and AC(1) perform better than that for  $\hat{\lambda}$ . Above a window length of  $T = 100$ , this difference is negligible, judging by the proposed metric. The difference may also be less pronounced when performing the same test on time series generated by models with non-linear drift or jump-noise, where the state-locality of our method can alleviate non-linear effects on the far ends of the state space. Therefore, in a large range of applications, the estimator  $\hat{\lambda}$  offers a statistically equally performant method of assessing CSD with the additional advantage of robustness with respect to time- and state-dependent noise, substantial advantages in higher-dimensional settings, as well as settings featuring non-linear drifts and jumps in the noise.

### 3. Details on the predator-prey model

The specific model introduced in the main text is a modification of the Truscott-Brindley model for ocean plankton populations originally introduced in [62]. Bengfort et al. [40] generalised the model by introducing the environmental parameter of fluid turbulence to the system and allowing higher powers in the mortality term of the predator population. The full system equations are given by

$$\xi \dot{p}(t) = rp(t) \left( 1 - \frac{p(t)}{K(\text{turb})} \right) - \frac{ap(t)^2}{h(\text{turb})^2 + p(t)^2} z(t)$$

$$\dot{z}(t) = \frac{ap(t)^2}{h(\text{turb})^2 + p(t)^2} z(t) - mz(t)^2$$

$$K(\text{turb}) = K_0 + c_K \cdot \text{turb}$$

$$h(\text{turb}) = \frac{h_0}{1 + c_h \cdot \text{turb}}$$

This system has been non-dimensionalised in order to reduce the number of parameters. However, to retrieve realistic values of population sizes in units of density,  $p$  and  $z$  merely need to be multiplied with constants  $p_0$  and  $z_0$ . The first term on the right-hand side of the prey population's evolution  $\dot{p}$  is the population growth rate as determined by the relationship between the current population size and the carrying capacity  $K$ . Below that capacity, the population grows and vice versa. The second term is the mortality rate of the prey population, which is simultaneously the growth rate of the predator population since it is assumed that all death in  $p$  and growth in  $z$  occurs through consumption of the former by the latter. The second term in the evolution of  $z$  in the second equation is the quadratic mortality term alluded to in the main text. This ultimately facilitates multiple stable states as opposed to the same model with a linear mortality term. The turbulence  $\text{turb} \in [0, 1]$  describes the normalised strength of spatial mixing in the ocean modelled by circular eddies. All parameter values but those for  $\xi$ ,  $c_K$  and  $c_h$  are adopted directly from [40] and can be found in Table SI along with a short description of their interpretation. The parameters  $c_K$  and  $c_h$  were increased by a factor of 2.2 each for the purposes of this study to facilitate a bigger range of stable prey populations in the large population regime. The fundamental nature of the model remains unaltered by this change. Lastly, as described in the main text, we introduced multiplicative noise terms commonly used in the relevant literature [42–45] to model environmental impacts on the growth and mortality rates of the two populations. This leads us to

TABLE SI. Parameter values used in the simulation of plankton populations following the model in equations (7).

parameter	value	description
$r$	1	growth rate factor of prey $P$
$a$	1/9	rate of the predator consuming the prey
$m$	0.0525	mortality rate of the predator
$\xi$	0.7	time scale separation between prey and predator evolutions
$h_0$	1/16	factor influencing maximal consumption at zero turbulence
$c_h$	0.88	linear relationship between turbulence and $h$
$K_0$	0.7	carrying capacity at zero turbulence
$c_K$	0.66	linear relationship between turbulence and $K$
$\sigma_P$	0.037	strength of noise coupling to $P$
$\sigma_Z$	0.01	strength of noise coupling to $Z$

the complete set of model equations:

$$\begin{aligned}
dP_t &= \xi^{-1} \left( rP_t \left( 1 - \frac{P_t}{K(\text{turb})} \right) - \frac{aP_t^2}{h(\text{turb})^2 + P_t^2} Z_t \right) dt \\
&\quad + \xi^{-1/2} \sigma_P P_t dW_t^P, \\
dZ_t &= \left( \frac{aP_t^2}{h(\text{turb})^2 + P_t^2} Z_t - mZ_t \right) dt + \sigma_Z Z_t dW_t^Z. \\
K(\text{turb}) &= K_0 + c_K \cdot \text{turb}, \\
h(\text{turb}) &= \frac{h_0}{1 + c_h \cdot \text{turb}} \\
\text{turb}(t) &= 1 - \frac{7}{10} \frac{t}{T}.
\end{aligned} \tag{7}$$

#### 4. Additional example of the multi-dimensional stability analysis

In the example of the two-dimensional predator-prey model in the main text, it was revealed by the analysis of the two-dimensional drift coefficient that the dynamics could also be well-represented by two uncoupled one-dimensional SDEs for the predator and prey population, respectively. As a result, the CSD analysis was also comprehensive after a reduction to the prey dimension. However, the dynamics of many systems cannot be reduced in such a way. This is especially relevant for systems exhibiting pronounced oscillations. Using the method outlined above, we therefore additionally assess the local stability of a system undergoing a subcritical Hopf bifurcation in normal form.

$$\begin{aligned}
d\mathbf{X}_t &= \begin{pmatrix} -\left( \mu(t) - (\mathbf{X}^{(1)})^2 - (\mathbf{X}^{(2)})^2 \right) \mathbf{X}^{(1)} - \omega \mathbf{X}^{(2)} \\ -\left( \mu(t) - (\mathbf{X}^{(1)})^2 - (\mathbf{X}^{(2)})^2 \right) \mathbf{X}^{(2)} + \omega \mathbf{X}^{(1)} \end{pmatrix} dt \\
&\quad + \varepsilon \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} d\mathbf{W}_t,
\end{aligned}$$

where  $\omega = 1$ ,  $\varepsilon = 0.01$ , and  $\mu(t)$  decreases linearly from 2 to 0.1 over the integration time of  $T = 1000$ . For  $\mu > 0$ , the origin is a stable fixed point with eigenvalues  $-\mu \pm i\omega$ . Furthermore, there is an unstable limit cycle with radius

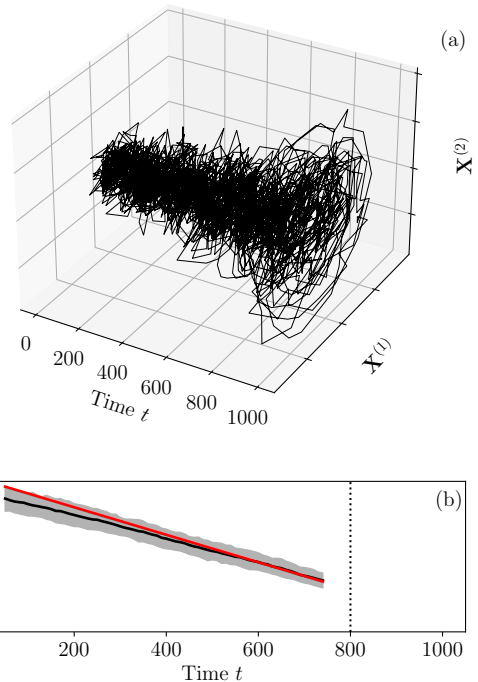


FIG. S3. Employing the outlined method for multidimensional stability analysis. (a) Time series data generated by a model undergoing a subcritical Hopf bifurcation in normal form. (b) Estimates of the stability indicators  $\lambda_k$ . The thick line represents the mean of the 200 sample estimations. Since the Jacobian matrices at each point in time have complex eigenvalues corresponding to the oscillatory dynamics, the eigenvalues are conjugates and coincide in their real part. The theoretical value for the negative of this real part is plotted in red.

$\sqrt{\mu}$  and perturbations from the origin decay in the form of spirals. At  $\mu = 0$ , the radius of the unstable limit cycle reaches zero, and the origin turns into an unstable fixed point. The data was sampled at time steps  $\Delta t = 0.1$



and analysed in windows of length  $T = 100$ . The results are presented in Fig. S3. The real parts of both eigenvalues are known to be  $-\mu(t)$ , and the estimations track this value relatively closely. A destabilisation of the equilibrium can clearly be made out. An additional insight

gained via the CSD assessment through the Langevin equation approach proposed here is that the local system exhibits oscillatory dynamics, as identified by the complex eigenvalues of the Jacobian matrix.

- 
- [1] T. Lenton, H. Held, E. Kriegler, J. W. Hall, W. Lucht, S. Rahmstorf, and H. J. Schellnhuber, Tipping elements in the Earth's climate system, *Proceedings of the National Academy of Sciences of the United States of America* **105**, 1786 (2008).
- [2] M. Scheffer, J. Bascompte, W. A. Brock, V. Brovkin, S. R. Carpenter, V. Dakos, H. Held, E. H. van Nes, M. Rietkerk, and G. Sugihara, Early-warning signals for critical transitions, *Nature* **461**, 53 (2009).
- [3] N. Boers, Observation-based early-warning signals for a collapse of the Atlantic Meridional Overturning Circulation, *Nature Climate Change* **11**, 680 (2021).
- [4] N. Boers, M. Ghil, and T. F. Stocker, Theoretical and paleoclimatic evidence for abrupt transitions in the earth system, *Environmental Research Letters* **17**, 093006 (2022).
- [5] R. Zwanzig, *Nonequilibrium Statistical Mechanics* (Oxford University Press, 2001).
- [6] K. Hasselmann, Stochastic climate models Part I. Theory, *Tellus* **28**, 473 (1976).
- [7] W. Horsthemke and R. Lefever, *Noise-Induced Transitions* (Springer Berlin, 1984).
- [8] C. Meisel, A. Klaus, C. Kuehn, and D. Plenz, Critical slowing down governs the transition to neuron spiking, *PLOS Computational Biology* **11**, e1004097 (2015).
- [9] C. Boettner and N. Boers, Critical slowing down in dynamical systems driven by nonstationary correlated noise, *Physical Review Research* **4**, 013230 (2022).
- [10] A. Morr, D. Kreher, and N. Boers, Red noise in continuous-time stochastic modelling, arXiv (2022).
- [11] P. Ashwin, S. Wieczorek, R. Vitolo, and P. Cox, Tipping points in open systems: bifurcation, noise-induced and rate-dependent examples in the climate system, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **370**, 1166 (2012).
- [12] V. Dakos, M. Scheffer, E. H. van Nes, V. Brovkin, V. Petoukhov, and H. Held, Slowing down as an early warning signal for abrupt climate change, *Proceedings of the National Academy of Sciences* **105**, 14308 (2008).
- [13] P. deMenocal, J. Ortiz, T. Guilderson, J. Adkins, M. Sarnthein, L. Baker, and M. Yarusinsky, Abrupt onset and termination of the African Humid Period: rapid climate responses to gradual insolation forcing, *Quaternary Science Reviews* **19**, 347 (2000).
- [14] N. Boers, N. Marwan, H. M. J. Barbosa, and J. Kurths, A deforestation-induced tipping point for the South American monsoon system, *Scientific Reports* **7**, 41489 (2017).
- [15] C. Kuehn, A mathematical framework for critical transitions: Bifurcations, fast-slow systems and stochastic dynamics, *Physica D: Nonlinear Phenomena* **240**, 1020 (2011).
- [16] H. Held and T. Kleinen, Detection of climate system bifurcations by degenerate fingerprinting, *Geophysical Research Letters* **31**, 10.1029/2004GL020972 (2004).
- [17] V. Dakos, Identifying best-indicator species for abrupt transitions in multispecies communities, *Ecological Indicators* **94**, 494 (2018).
- [18] C. Kuehn, A mathematical framework for critical transitions: Normal forms, variance and applications, *Journal of Nonlinear Science* **23**, 457 (2013).
- [19] T. M. Lenton, V. N. Livina, V. Dakos, E. H. van Nes, and M. Scheffer, Early warning of climate tipping points from critical slowing down: comparing methods to improve robustness, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **370**, 1185 (2012).
- [20] G. E. Uhlenbeck and L. S. Ornstein, On the theory of the brownian motion, *Phys. Rev.* **36**, 823 (1930).
- [21] E. Weinans, R. Quax, E. H. van Nes, and I. A. v. d. Leemput, Evaluating the performance of multivariate indicators of resilience loss, *Scientific Reports* **11**, 9148 (2021).
- [22] V. Dakos, E. H. van Nes, P. D'Odorico, and M. Scheffer, Robustness of variance and autocorrelation as indicators of critical slowing down, *Ecology* **93**, 264 (2012).
- [23] J. J. Clarke, C. Huntingford, P. D. L. Ritchie, and P. M. Cox, Seeking more robust early warning signals for climate tipping points: the ratio of spectra method (ROSA), *Environmental Research Letters* **18**, 035006 (2023).
- [24] D. Lamouroux and K. Lehnertz, Kernel-based regression of drift and diffusion coefficients of stochastic processes, *Physics Letters A* **373**, 3507 (2009).
- [25] L. Rydin Gorjão and F. Meirinhos, `kramersmoyal`: Kramers-moyal coefficients for stochastic processes, *Journal of Open Source Software* **4**, 1693 (2019).
- [26] T. Wand, M. Heßler, and O. Kamps, Memory effects, multiple time scales and local stability in langevin models of the s&p500 market correlation (2023), arXiv:2307.12744 [q-fin.ST].
- [27] H. Risken and T. Frank, *The Fokker-Planck equation*, 2nd ed. (Springer-Verlag, Berlin, Heidelberg, 1996).
- [28] N. G. van Kampen, Chapter VIII - The Fokker-Planck equation, in *Stochastic Processes in Physics and Chemistry*, North-Holland Personal Library (Elsevier, 2007) 3rd ed., pp. 193-218.
- [29] C. Gardiner, *Stochastic Methods: A Handbook for the Natural and Social Sciences*, 4th ed. (Springer-Verlag Berlin Heidelberg, 2009).
- [30] C. Honisch and R. Friedrich, Estimation of Kramers-Moyal coefficients at low sampling rates, *Physical Review E* **83**, 066701 (2011).
- [31] M. R. R. Tabar, *Analysis and Data-Based Reconstruction of Complex Nonlinear Dynamical Systems*, 1st ed. (Springer International Publishing, 2019).
- [32] F. Hassanibesheli, N. Boers, and J. Kurths, Reconstructing complex system dynamics from time series: a method comparison, *New Journal of Physics* **22**, 073053 (2020).
- [33] K. Riechers, L. Rydin Gorjão, F. Hassanibesheli, P. G. Lind, D. Witthaut, and N. Boers, Stable stadial and in-

- terstadial states of the last glacial's climate identified in a combined stable water isotope and dust record from greenland, *Earth System Dynamics* **14**, 593 (2023).
- [34] L. Rydin Gorjão, J. Heysel, K. Lehnertz, and M. R. R. Tabar, Analysis and data-driven reconstruction of bivariate jump-diffusion processes, *Physical Review E* **100**, 062127 (2019).
- [35] B. M. S. Arani, S. R. Carpenter, L. Lahti, E. H. van Nes, and M. Scheffer, Exit time as a measure of ecological resilience, *Science* **372**, eaay4895 (2021).
- [36] H. A. Kramers, Brownian motion in a field of force and the diffusion model of chemical reactions, *Physica* **7**, 284 (1940).
- [37] V. Guttal and C. Jayaprakash, Impact of noise on bistable ecological systems, *Ecological Modelling* **201**, 420 (2007).
- [38] E. M. Fischer and C. Schär, Future changes in daily summer temperature variability: driving processes and role for temperature extremes, *Climate Dynamics* **33**, 917 (2009).
- [39] R. J. Shiller, Causes of changing financial market volatility (1988) pp. 1–28, <https://www.kansascityfed.org/Jackson%20Hole/documents/3811/1988-S88SHILL.pdf>.
- [40] M. Bengfort, U. Feudel, F. M. Hilker, and H. Malchow, Plankton blooms and patchiness generated by heterogeneous physical environments, *Ecological Complexity* **20**, 185 (2014).
- [41] M. Sieber, H. Malchow, and L. Schimansky-Geier, Constructive effects of environmental noise in an excitable prey–predator plankton system with infected prey, *Ecological Complexity* **4**, 223 (2007).
- [42] R. M. May, Stability in randomly fluctuating versus deterministic environments, *The American Naturalist* **107**, 621 (1973).
- [43] G. P. Samanta, A stochastic two species competition model: nonequilibrium fluctuation and stability, *International Journal of Stochastic Analysis* **2011**, 489386 (2011).
- [44] A. Das and G. P. Samanta, Influence of environmental noises on a prey–predator species with predator-dependent carrying capacity in alpine meadow ecosystem, *Mathematics and Computers in Simulation* **190**, 1294 (2021).
- [45] M. Belabbas, A. Ouahab, and F. Soua, Rich dynamics in a stochastic predator-prey model with protection zone for the prey and multiplicative noise applied on both species, *Nonlinear Dynamics* **106**, 2761 (2021).
- [46] N. Boers, Early-warning signals for Dansgaard-Oeschger events in a high-resolution ice core record, *Nature Communications* **9**, 2556 (2018).
- [47] T. M. Bury, C. T. Bauch, and M. Anand, Detecting and distinguishing tipping points using spectral early warning signals, *Journal of The Royal Society Interface* **17**, 20200482 (2020).
- [48] P. Hines, E. Cotilla-Sanchez, and S. Blumsack, Topological models and critical slowing down: Two approaches to power system blackout risk analysis, in *2011 44th Hawaii International Conference on System Sciences* (2011) pp. 1–10.
- [49] E. Cotilla-Sanchez, P. D. H. Hines, and C. M. Danforth, Predicting critical transitions from time series synchrophasor data, *IEEE Transactions on Smart Grid* **3**, 1832 (2012).
- [50] H. Ren and D. Watts, Early warning signals for critical transitions in power systems, *Electric Power Systems Research* **124**, 173 (2015).
- [51] V. Guttal, S. Raghavendra, N. Goel, and Q. Hoarau, Lack of critical slowing down suggests that financial meltdowns are not critical transitions, yet rising variability could signal systemic risk, *PLOS ONE* **11**, e0144198 (2016).
- [52] H. Gatfaoui, I. Nagot, and P. de Peretti, 3 - are critical slowing down indicators useful to detect financial crises?, in *Systemic Risk Tomography*, edited by M. Billo, L. Pelizzon, and R. Savona (Elsevier, 2017) pp. 73–93.
- [53] C. Diks, C. Hommes, and J. Wang, Critical slowing down as an early warning signal for financial crises?, *Empirical Economics* **57**, 1201 (2019).
- [54] A. Levermann, J. Schewe, V. Petoukhov, and H. Held, Basic mechanism for abrupt monsoon transitions, *Proceedings of the National Academy of Sciences* **106**, 20572 (2009).
- [55] J. Schewe, A. Levermann, and H. Cheng, A critical humidity threshold for monsoon transitions, *Climate of the Past* **8**, 535 (2012).
- [56] C. A. Boulton, T. M. Lenton, and N. Boers, Pronounced loss of Amazon rainforest resilience since the early 2000s, *Nature Climate Change* **12**, 271 (2022).
- [57] T. Smith, D. Traxl, and N. Boers, Empirical evidence for recent global shifts in vegetation resilience, *Nature Climate Change* **12**, 477 (2022).
- [58] G. Forzieri, V. Dakos, N. G. McDowell, A. Ramdane, and A. Cescatti, Emerging signals of declining forest resilience under climate change, *Nature* **608**, 534 (2022).
- [59] T. Smith and N. Boers, Global vegetation resilience linked to water availability and variability, *Nature Communications* **14**, 498 (2023).
- [60] A. Levermann and R. Winkelmann, A simple equation for the melt elevation feedback of ice sheets, *The Cryosphere* **10**, 1799 (2016).
- [61] N. Boers and M. Rypdal, Critical slowing down suggests that the western greenland ice sheet is close to a tipping point, *Proceedings of the National Academy of Sciences* **118**, e2024192118 (2021).
- [62] J. E. Truscott and J. Brindley, Ocean plankton populations as excitable media, *Bulletin of Mathematical Biology* **56**, 981 (1994).